CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Research Papers in Economics

On Testing for Unit Roots and the Initial Observation

David I. Harvey
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD, UK
dave.harvey@nottingham.ac.uk

Stephen J. Leybourne
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD, UK
steve.leybourne@nottingham.ac.uk

We consider the power of unit root tests for different deviations of the initial observation from the deterministic component of the series. Following recent work highlighting the relative power performance of extant tests, we propose a new procedure based on a data-dependent weighted average of the standard Dickey-Fuller and Elliott-Rothenberg-Stock tests, with the weight determined by an estimate of the initial observation's deviation from the deterministics. Simulation of the new test's power reveals very good performance across different magnitudes of the initial condition. The procedure's value is further highlighted by application to US producer price inflation.

KEY WORDS: Dickey-Fuller test; GLS detrending; Power comparison; Weighted average.

The issue of whether an economic time series is best characterised by either a unit root or a stationary process has assumed great importance in both the theoretical and applied time series econometrics literature. As a consequence, tests of the null hypothesis that a series is integrated of order one, I(1), against the alternative hypothesis that it is integrated of order zero, I(0), have received much attention. In a recent important paper, Müller and Elliott (2003) extend the literature on these testing procedures by highlighting the dependence of the power of all unit root tests on the deviation of the initial observation of the series from its underlying determinstic component. These authors show that the well-known augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) has power that increases with the magnitude of this initial deviation, although for small initial conditions, power is dominated by other available procedures. On the other hand, the test proposed by Elliott, Rothenberg and Stock (ERS, 1996), based on GLS-detrending, is optimal when the initial deviation is zero, but the test's power shrinks to zero as the deviation of the initial observation from the deterministic component becomes large.

In this paper, we consider an alternative approach which attempts to try and capture the attractive power performance of both the ERS test for small initial conditions and the ADF test for large initial conditions. The technique we propose is to use a data-dependent weighted average of the ADF and ERS test statistics, with the weight determined by an estimate of the deviation of the initial observation from the underlying deterministics. This procedure is shown to have reliable power performance across different initial observations, almost dominating ADF while retaining much of the ERS power when the initial deviation is small.

The paper is structured as follows: Section 1 reviews in detail the power of the ADF and ERS tests for a range of initial conditions, using Monte Carlo simulation. Section 2 introduces the new weighted average test, provides its asymptotic distribution and investigates its power performance, again by simulation. The value of the new procedure is further highlighted by application of the extant and new tests to US producer price inflation in Section 3. Section 4 concludes the paper.

1. POWERS OF STANDARD TESTS

Consider the following data generating process (cf. Müller and Elliott, 2003):

$$
\begin{aligned}
y_t &= d_t + w_t \qquad t = 1, \ldots, T \\
w_t &= \rho w_{t-1} + \nu_t \quad t = 2, \ldots, T \\
w_1 &= \xi
\end{aligned} \tag{1}
$$

where d_t denotes the deterministic component of the series, and ν_t is a stationary process. In this paper we consider the two standard forms of deterministics: the mean case, where $d_t = \mu$, and the trend case, where $d_t = \mu + \beta t$. Following Müller and Elliott (2003), let $\xi = \alpha\sigma_w$, where σ_w^2 is the unconditional variance of w_t (defined over $t = 2, \ldots, T$), so that α represents the magnitude of the initial observation of w_t relative to the process standard deviation.

In this framework, Müller and Elliott (2003) note that, under the null hypothesis, standard unit root tests such as ADF and ERS will be invariant to the initial observation. Thus, the standard null limit distributions of these test statistics are valid irrespective of the magnitude of α, and the test procedures have correct size. It is under the alternative hypothesis that the initial observation has an effect, and the powers of tests vary according to α and the testing approach employed.

Müller and Elliott (2002, 2003) examine the powers of several extant unit root tests for a range of α values. Their results show that when α is zero or small, the t-ratio versions of the GLS-detrended tests proposed by ERS achieve the greatest power, while for larger α values, standard t-ratio ADF tests perform best. Focusing on these two popular approaches, we conducted a set of Monte Carlo simulation experiments to investigate the tests' powers in detail. We simulated both finite sample and asymptotic behaviour for the mean and trend cases under the local-to-unity alternative hypothesis, with $\rho = 1 + c/T$, $c = -5, -10, -15$ and $T = 100, 200, \infty$, for tests with 5% nominal size (the ERS tests were implemented with the recommended $\bar{c}$ values of -7 and -13.5 for the mean and trend cases respectively). The finite sample experiments were conducted with $\nu_t \sim \text{IIN}(0, 1)$ in (1), and with no additional lag augmentation in the test statistics. The limiting distributions are given in Müller and Elliott (2003), and were simulated by approximating the Wiener processes using $\text{IIN}(0, 1)$ random deviates, with the integrals approximated by normalized sums of 1,000 steps. Critical values were first obtained in each case using 50,000 replications, and the subsequent power experiments were conducted using 20,000 replications. All calculations were performed in GAUSS. Figures 1–3 provide the results of these experiments (ADF and ERS tests are denoted by $\hat{\tau}^i_{\text{ADF}}$ and $\hat{\tau}^i_{\text{ERS}}$

respectively, with $i = \mu, \tau$ corresponding to the mean and trend cases respectively), and, where comparable, these concur with those of Müller and Elliott (2002, 2003).

The ERS tests have greater power than the ADF procedures for α values approximately less than one in the mean case, and approximately less than 1.5 in the trend case, with the degree of power dominance increasing as α shrinks towards zero. Reverse rankings are obtained for greater α magnitudes: as α increases, so does ADF test power, while the power of ERS tests converges to zero. This pattern of behaviour arises since ADF tests implicitly place a large weight on extreme initial observations, in contrast to the tests of ERS which are derived under an assumption of ξ bounded in probability, attributing more weight to moderate deviations of the initial observation from the underlying process.

2. A WEIGHTED AVERAGE TEST

Given the results of the previous section, it is worthwhile considering whether a different testing procedure can be adopted which capitalises on the power of ERS tests when α is small, but also achieves the power gains of ADF tests when α is large. One possibility is to consider a simple weighted average of the two tests, i.e.

$$\hat{\tau}^i_{\mathrm{AV}} = \lambda \hat{\tau}^i_{\mathrm{ADF}} + (1-\lambda)\hat{\tau}^i_{\mathrm{ERS}}, \qquad 0 \le \lambda \le 1, \qquad i = \mu, \tau \tag{2}$$

Within this framework, the next issue concerns determination of the weight λ. One approach would be to use an exogenously chosen weight based on simulation results under the alternative hypothesis. In unreported simulations, we investigated this possibility for a grid of values for λ, and found that $\lambda = 0.75$ yielded the most appealing powers for different values of α. The powers of this procedure for the representative mean case of $T = 100$, $c = -10$ are given in Figure 4. The simulation experiments were identical to those of the previous section, and the weighted average test is denoted by $\hat{\tau}^\mu_{\mathrm{AV}}(0.75)$. It can be seen that this fixed weight approach has some value: for small values of α, the test outperforms ADF, achieving approximately half the gains offered by ERS, while for larger α values, the test shares the desirable ADF property of increasing power with α. Power never falls below 36%, and for $1 \le \alpha \le 1.7$, power is greater than for either of the constituent tests. The procedure is not entirely satisfactory, however, since for the majority of α values, the test's power is substantially outperformed by either the ADF or ERS approaches.

A natural generalisation of the weighted average approach, which might be expected to offer more appealing power performance, is to consider a data-dependent estimate of λ, rather than an exogenous choice. Since the behaviour of the unit root tests under the alternative hypothesis depend on the magnitude of α, it makes sense to base the estimated weight $\hat{\lambda}$ on an estimate of α. Given that $\alpha = \xi/\sigma_w$, the obvious estimator is:

$$\hat{\alpha} = (y_1 - \hat{d}_1)/\hat{\sigma}_w \tag{3}$$

where $\hat{d}_t$ and $\hat{\sigma}_w^2 = \hat{V}(w_t)$ are obtained from ordinary least squares estimation of the regression:

$$y_t = d_t + \varepsilon_t, \qquad t = 2, \ldots, T \tag{4}$$

with $\hat{d}_t$ denoting the fitted values and $\hat{\sigma}_w^2 = (T-1)^{-1} \sum_{t=2}^{T} \hat{\varepsilon}_t^2$.

The value of $\hat{\alpha}$ obtained from (3) must then be transformed into a weight $\hat{\lambda}$ on the interval $[0, 1]$ so that the weighted average statistic can be constructed. The method we propose is to apply the logistic smooth transition function (as used by, for example, Granger and Teräsvirta (1993) and Lin and Teräsvirta (1994) in the context of modelling structural change) to $\hat{\alpha}$, i.e.

$$\hat{\lambda} = [1 + \exp\{-v(\hat{\alpha} - m)\}]^{-1} \tag{5}$$

This function ensures that for small α, the weight in the linear combination will be zero, while for large α, the weight will be one, as desired. The mid-point parameter m determines the location in $\hat{\alpha}$ space where the balance of weight switches from ERS to ADF, while the velocity v controls the rate at which the weight moves from zero to one as $\hat{\alpha}$ increases (when $v \to \infty$, $\hat{\lambda}$ switches instantaneously from zero to one at $\hat{\alpha} = m$). The test resulting from this procedure is then given by:

$$\hat{\tau}_{\mathrm{AV}}^{i}(\hat{\alpha}; v, m) = \hat{\lambda}\hat{\tau}_{\mathrm{ADF}}^{i} + (1 - \hat{\lambda})\hat{\tau}_{\mathrm{ERS}}^{i}, \qquad i = \mu, \tau \tag{6}$$

The asymptotic distribution of the new weighted average statistic under the null ($\rho = 1$) and local alternative ($\rho = 1 + c/T$, $c < 0$) hypotheses can be obtained by application of the continuous mapping theorem (CMT) to (5) and (6), once the limiting distributions of $\hat{\tau}_{\mathrm{ADF}}^{i}$ and $\hat{\tau}_{\mathrm{ERS}}^{i}$ (assuming these tests are appropriately augmented to account for autocorrelation), and $\hat{\alpha}$ have been established. The first two of these required results are provided in Müller and Elliott (2003), and the last is derived in the Appendix.

Under Assumption 1 of the Appendix, the new test's asymptotic distributions are given by:

$$\hat{\tau}_{\mathrm{AV}}^{i}(\hat{\alpha}; v, m) \Rightarrow \lambda_c(v,m)\tau_{c,\mathrm{ADF}}^{i} + [1-\lambda_c(v,m)]\tau_{c,\mathrm{ERS}}^{i}, \qquad i = \mu, \tau \tag{7}$$

where '$\Rightarrow$' denotes weak convergence in distribution, $\lambda_c(v,m) = [1+\exp\{-v(A_c^i - m)\}]^{-1}$ and

$$\begin{aligned}
A_c^{\mu} &= \frac{-\int_0^1 K_c(r)dr}{\sqrt{\int_0^1 K_c^{\mu}(r)^2 dr}}, & A_c^{\tau} &= \frac{-\int_0^1 K_c(r)dr + 6\int_0^1 (r-\frac{1}{2})K_c(r)dr}{\sqrt{\int_0^1 K_c^{\tau}(r)^2 dr}} \\
\tau_{c,\mathrm{ADF}}^{\mu} &= \frac{K_c^{\mu}(1)^2 - K_c^{\mu}(0)^2 - 1}{2\sqrt{\int_0^1 K_c^{\mu}(r)^2 dr}}, & \tau_{c,\mathrm{ADF}}^{\tau} &= \frac{K_c^{\tau}(1)^2 - K_c^{\tau}(0)^2 - 1}{2\sqrt{\int_0^1 K_c^{\tau}(r)^2 dr}} \\
\tau_{c,\mathrm{ERS}}^{\mu} &= \frac{K_c(1)^2 - K_c(0)^2 - 1}{2\sqrt{\int_0^1 K_c(r)^2 dr}}, & \tau_{c,\mathrm{ERS}}^{\tau} &= \frac{K_c^{\tau,\bar{c}}(1)^2 - K_c^{\tau,\bar{c}}(0)^2 - 1}{2\sqrt{\int_0^1 K_c^{\tau,\bar{c}}(r)^2 dr}}
\end{aligned}$$

with

$$\begin{aligned}
K_c(r) &= \begin{cases} W(r) & c = 0 \\ \alpha(e^{rc}-1)(-2c)^{-1/2} + \int_0^r e^{(r-s)c} dW(s) & c < 0 \end{cases} \\
K_c^{\mu}(r) &= K_c(r) - \int_0^1 K_c(s)ds \\
K_c^{\tau}(r) &= K_c^{\mu}(r) - 12(r-\tfrac{1}{2})\int_0^1 (s-\tfrac{1}{2})K_c(s)ds \\
K_c^{\tau,\bar{c}}(r) &= K_c(r) - r(1-\bar{c})(1-\bar{c}+\tfrac{\bar{c}^2}{3})^{-1}K_c(1) + 3(1-r)\int_0^1 sK_c(s)ds
\end{aligned}$$

Under an alternative hypothesis where $\rho < 1$ is fixed, standard results show that $\hat{\mu}$ and $\hat{\beta}$ in $\hat{d}_t$ consistently estimate μ and β respectively, and also that $\hat{\sigma}_w \xrightarrow{p} \sigma_w$; thus $\hat{\alpha}$ is a consistent estimator of α as desired.

Now the critical values and power performance of the test will depend on the choice of the smooth transition parameters v and m. In order to determine appropriate values for these terms, we simulated the powers of $\hat{\tau}_{\mathrm{AV}}^{\mu}(\hat{\alpha}; v, m)$ and $\hat{\tau}_{\mathrm{AV}}^{\tau}(\hat{\alpha}; v, m)$ for a grid of v and m values, using experiments identical to those described in Section 2. Results for the mean case with $T = 100$ and $c = -10$ are reported in Table 1, allowing the impact for test power of changing the velocity and mid-point of the transition to be seen. Our preference is for parameters that yield a test whose minimum power is decent, while also achieving high power when $\alpha = 0$ and when α is large. Balancing these considerations, we concluded that the best row of Table 1 corresponds to a transition velocity of 0.75, and a mid-point of 1.25. This mid-point choice is also intuitively appealing since $\alpha = 1.25$

is approximately where ADF power starts to exceed ERS power, as seen in Figures 1–3. Simulations for a finer grid of transition parameter values revealed a little improvement by reducing the velocity slightly, and our recommended parameter choices for $\hat{\tau}^{\mu}_{\mathrm{AV}}(\hat{\alpha}; v, m)$ are $v = 0.73$, $m = 1.25$. A similar analysis for the trend case found that these values were also appropriate for $\hat{\tau}^{\tau}_{\mathrm{AV}}(\hat{\alpha}; v, m)$.

Finite sample and asymptotic critical values for the weighted average tests at conventional significance levels are provided in Table 2. As before, these were obtained by simulation using the method outlined in Section 1. The powers of our recommended tests were simulated for the full range of sample sizes and values of c considered earlier in the paper, and the results are shown in Figures 1–3. These graphs clearly highlight the value of the new procedure: the weighted average test is close to dominating the ADF approach, while achieving much of the small α power gains of ERS, and there is always a region of α values where the new test outperforms both of the standard approaches.

Examining a relevant case such as $T = 100$, $c = -10$, power is always greater than 40% (as opposed to 31% and 0% for ADF and ERS respectively), is close to 60% for $\alpha = 0$ (compared to 31% and 73% for the standard tests), and shares the ADF power of 70% for $\alpha = 4$ (in contrast to ERS zero power). Indeed the power of the weighted average test is not far off the envelope of the powers of ADF and ERS together, exceeding it for moderate α, and only losing out by around 13% for the smallest values of α. Although the powers involved vary across the different T and c values considered, and from the mean to the trend case, the overall patterns of behaviour under the alternative described in this representative case are observed generally.

Figure 4 also shows the comparison of this new test against the fixed weight test $\hat{\tau}^{\mu}_{\mathrm{AV}}(0.75)$. The data-dependent smooth transition approach clearly dominates the exogenously chosen weight method, with greater power for all α values, particularly at the extremes.

The proposed procedure relies on knowledge of the sign of α, and the analysis thus far has implicitly assumed $\alpha \geq 0$; if it is known that $\alpha < 0$, the appropriate modification is simply to replace $\hat{\alpha}$ in (5) with $-\hat{\alpha}$ (the critical values are unaffected by such a transformation). This assumption of the sign of α being known is likely to most admissible when α is relatively large. However, even when α is relatively small and some doubt exists concerning its true sign, we still recommend maintaining the "known sign" assumption. Our justification for this is on grounds of power. If the known sign assumption is dismissed, the obvious alternative is to employ a new procedure with $|\hat{\alpha}|$ replacing $\hat{\alpha}$ in (5), thereby permitting the analyst to be truly agnostic with regard to α's sign. Such a modification clearly alters the critical values (they will be larger in absolute value), and impacts the power of the test. Using the now-familiar simulation analysis, we estimated the power of this procedure (denoted $\hat{\tau}^{i}_{\mathrm{AV}}(|\hat{\alpha}|; 0.73, 1.25)$), for a range of positive and negative values of α in the representative mean case of $T = 100$, $c = -10$. Figure 5 contains the results of this experiment, along with the corresponding powers for $\hat{\tau}^{\mu}_{\mathrm{ADF}}$, $\hat{\tau}^{\mu}_{\mathrm{ERS}}$ and $\hat{\tau}^{\mu}_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$ (always assuming $\alpha > 0$). For positive values of α, the power of the "absolute" procedure loses approximately 10% power relative to the approach assuming $\alpha > 0$ in each case. Moreover, for small negative values of α, the known sign version of the test outpeforms the absolute version, even though the wrong sign is being assumed in these cases. It is only as α becomes more substantially negative that the absolute version achieves gains, but as α decreases away from zero, the true negative sign of α will rapidly become obvious, thereby justifying the known sign assumption again, only now using $-\hat{\alpha}$. Therefore, although there is a small region where the power of the absolute approach gains relative to the known sign procedure, we consider such gains to be far outweighed by the overall power losses that the former technique exhibits for α values in general.

3. EMPIRICAL APPLICATION

In this section, we assess the behaviour of the newly proposed and standard unit root tests when applied to producer price inflation in the US. The data are first differences of the logarithms of the US producer price index, using monthly observations from 1973:1–2003:3, obtained from the Economagic website (www.economagic.com). A plot of this time series is given in Figure 6. Visual inspection clearly suggests that the series is mean-reverting, thus we would expect, *a priori*, unit root tests to reject the null hypothesis.

In order to examine how the different tests behave under different initial conditions, we applied the $\hat{\tau}^{\mu}_{\mathrm{ADF}}$, $\hat{\tau}^{\mu}_{\mathrm{ERS}}$ and $\hat{\tau}^{\mu}_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$ tests to the series using thirty different starting points. The start date varies from 1973:1–1975:6, with the sample size ranging from 334–363 observations. This period was chosen because it includes a number of both relatively large and relatively small initial observations, so the effect of the beginning of the series on the tests will be highlighted. Since the initial observations over this period are predominantly above mean, we generally assumed knowledge of $\alpha > 0$; the exceptions were three start dates (1973:7, 1973:9, 1973:10) where the initial observation is clearly below mean—in these cases, we assumed knowledge of $\alpha < 0$. In order to allow for additional autocorrelation in the series, eleven difference lags were included in the ADF and ERS regressions to admit AR(12) dynamics. Asymptotic critical values of the tests were employed in each case. The results of this empirical application are provided in Table 3.

The ADF test rejects the unit root null in favour of I(0) behaviour for the first fifteen of the thirty starting values, but not the latter fifteen. In general, the deviations of the initial observations from the mean are much larger in absolute value for the first half of the starting values considered than for the second half; this can be observed from Figure 6 and also from the reported values of $\hat{\alpha}$ in Table 3. The ERS test, in contrast, only rejects the unit root null when the initial observation is small (more specifcally when $|\hat{\alpha}| < 0.7$), with only seven rejections out of the thirty series analysed. The pattern of rejections for the ADF and ERS tests are broadly consistent with the results from the simulation experiments discussed in Section 1, with the magnitude of the initial observation playing a large role in the power performance of the tests.

The new weighted average test rejects the null hypothesis much more frequently than either of the standard tests: rejections in favour of the I(0) alternative are obtained for all but six of the thirty starting values. The new approach completely dominates ADF and ERS in this application, in that whenever either or both of the ADF and ERS tests reject, the weighted average test also rejects, plus there are four further cases where neither ADF nor ERS reject, but the new procedure does. These findings are consistent with the simulation results obtained in the previous section, and clearly highlight the benefits of employing the new test. The weighted average approach is less sensitive to the initial condition than its rivals, and provides a reliable method for testing for a unit root, with decent power obtained regardless of the magnitude of the initial observation relative to the underlying process.

4. CONCLUSION

In practical applications, it is common to find series where the initial observation is small relative to the series mean or trend, and also series where the initial condition is large. Given that the power of currently available tests varies considerably with the magnitude of the initial condition, it is worthwhile to have available a unit root testing procedure whose power is more robust to the deviation of the initial observation from the underlying deterministic component. One would not, however, wish to sacrifice too much power in order to obtain such robustness. The weighted average test proposed in this paper achieves good power performance for all initial observation magnitudes, unlike the standard ADF and ERS tests currently used in the literature; the new procedure shares the power advantages of ADF for large initial deviations, while retaining most of the ERS power gains when the initial condition is small. Thus, we would strongly recommend use of the new test in empirical applications, since it broadly achieves the desired robustness to the initial condition without large power sacrifices relative to available alternatives.

APPENDIX: ASYMPTOTIC DISTRIBUTION OF $\hat{\alpha}$

Since $\hat{\alpha}$ of (3) is invariant to μ and β in d_t, we can let $\mu = \beta = 0$ without loss of generality, i.e. $y_t = w_t$. Further, under both the mean and trend cases, the regressor d_t in (4) contains at least a mean component, thus $\hat{\alpha}$ can equivalently be obtained from the transformed regression:

$$w_t^* = d_t + \varepsilon_t^*, \qquad t = 2, \dots, T$$

where $w_t^* = w_t - w_1$. Denoting the fitted values obtained from the regression by $\hat{d}_t^*$, an alternative expression for $\hat{\alpha}$ is given by

$$\hat{\alpha} = (w_1^* - \hat{d}_1^*)/\hat{\sigma}_w^* = -\hat{d}_1^*/\hat{\sigma}_w^*$$

where ${\hat{\sigma}_w^*}^2 = (T-1)^{-1} \sum_{t=2}^T \hat{\varepsilon}_t^{*2}$.

Decomposing w_t^* yields

$$w_t^* = \tilde{w}_t + (\rho^{t-1} - 1) w_1$$

where $\tilde{w}_t = \rho \tilde{w}_t + \nu_t$ with $\tilde{w}_1 = 0$.

Assumption 1: The stationary sequence ν_t has a strictly positive spectral density function; it has a moving average representation $\nu_t = \sum_{j=0}^{\infty} \delta_j \eta_{t-j}$ where the η_t are identically and independently distributed random variables with finite second and fourth moments, and $\sum_{j=0}^{\infty} j|\delta_j| < \infty$.

Under the local alternative hypothesis and Assumption 1, $\sigma_w^2 = \omega^2 T/(-2c) + o(T)$, yielding

$$\begin{aligned} T^{-1/2} w^*_{[rT]} &= T^{-1/2}\tilde{w}_{[rT]} + (\rho^{[rT]-1} - 1)\alpha\sqrt{\omega^2/(-2c) + o(1)} \\ &\Rightarrow \omega \textstyle\int_0^r e^{(r-s)c} dW(s) + (e^{rc} - 1)\alpha\omega(-2c)^{-1/2} \end{aligned}$$

where ω^2 is the long-run variance of ν_t and $W(r)$ is a standard Wiener process, using the standard result of Phillips (1987) and the fact that $\rho^{[rT]-1} \rightarrow e^{rc}$. The limiting behaviour under the null hypothesis can be obtained by letting $c \rightarrow 0$, with the above result reducing to $T^{-1/2} w^*_{[rT]} \Rightarrow \omega W(r)$. Summarising we have

$$T^{-1/2} w^*_{[rT]} \Rightarrow \omega K_c(r)$$

where

$$K_c(r) = \begin{cases} W(r) & c = 0 \\ \alpha(e^{rc} - 1)(-2c)^{-1/2} + \int_0^r e^{(r-s)c} dW(s) & c < 0 \end{cases}$$

as in Müller and Elliott (2003).

In the mean case, applying the CMT to the above result gives:

$$\begin{aligned} \hat{d}^*_t &= (T-1)^{-1} \textstyle\sum_{s=2}^T w^*_s \\ T^{-1/2}\hat{d}^*_1 &\Rightarrow \omega \textstyle\int_0^1 K_c(r) dr \end{aligned}$$

and

$$\begin{aligned} \hat{\varepsilon}^*_t &= w^*_t - (T-1)^{-1} \textstyle\sum_{s=2}^T w^*_s \\ T^{-1/2}\hat{\varepsilon}^*_{[rT]} &\Rightarrow \omega \left[K_c(r) - \textstyle\int_0^1 K_c(s) ds\right] \equiv \omega K_c^\mu(r) \end{aligned}$$

allowing the limiting distribution of $\hat{\alpha}$ in the mean case to be obtained using the CMT:

$$\begin{aligned} \hat{\alpha} &= -T^{-1/2}\hat{d}^*_1 / \sqrt{T^{-1}\hat{\sigma}^{*\,2}_w} \\ &\Rightarrow \frac{-\int_0^1 K_c(r) dr}{\sqrt{\int_0^1 K_c^\mu(r)^2 dr}} \equiv A_c^\mu \end{aligned}$$

In the trend case we have:

$$\hat{d}^*_t = (T-1)^{-1} \textstyle\sum_{s=2}^T w^*_s + \hat{\beta}^* \left[t - (T-1)^{-1} \textstyle\sum_{s=2}^T s\right]$$

where

$$\begin{aligned} \hat{\beta}^* &= \left[\textstyle\sum_{s=2}^T s w^*_s - (T-1)^{-1} \sum_{s=2}^T w^*_s \sum_{s=2}^T s\right] / \left[\textstyle\sum_{s=2}^T s^2 - (T-1)^{-1} (\sum_{s=2}^T s)^2\right] \\ T^{1/2}\hat{\beta}^* &\Rightarrow 12\omega \textstyle\int_0^1 (r - \frac{1}{2}) K_c(r) dr \end{aligned}$$

giving

$$T^{-1/2}\hat{d}^*_1 \Rightarrow \omega \left[\textstyle\int_0^1 K_c(r) dr - 6 \int_0^1 (r - \frac{1}{2}) K_c(r) dr\right]$$

and

$$\begin{aligned}
\hat{\varepsilon}_t^* &= w_t^* - (T-1)^{-1}\textstyle\sum_{s=2}^{T} w_s^* - \hat{\beta}^* \left[t - (T-1)^{-1}\textstyle\sum_{s=2}^{T} s\right] \\
T^{-1/2}\hat{\varepsilon}_{[rT]}^* &\Rightarrow \omega \left[K_c^\mu(r) - 12(r - \tfrac{1}{2}) \textstyle\int_0^1 (s - \tfrac{1}{2}) K_c(s) ds\right] \equiv \omega K_c^\tau(r)
\end{aligned}$$

The limiting distribution of $\hat{\alpha}$ in the trend case is then:

$$\begin{aligned}
\hat{\alpha} &= -T^{-1/2}\hat{d}_1^* / \sqrt{T^{-1}\hat{\sigma}_w^{*\,2}} \\
&\Rightarrow \frac{-\int_0^1 K_c(r)dr + 6\int_0^1 (r - \frac{1}{2}) K_c(r) dr}{\sqrt{\int_0^1 K_c^\tau(r)^2 dr}} \equiv A_c^\tau
\end{aligned}$$

REFERENCES

Dickey, D. A., and Fuller, W. A. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

Elliott, G. (1999), "Efficient Tests for a Unit Root when the Initial Observation is Drawn from its Unconditional Distribution," *International Economic Review*, 40, 767–783.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Granger, C. W. J., and Teräsvirta, T. (1993), *Modelling Nonlinear Economic Relationships*, Oxford: Oxford University Press.

Lin, C. J., and Teräsvirta, T. (1994), "Testing the Constancy of Regression Parameters Against Continuous Structural Change," *Journal of Econometrics*, 62, 211–228.

Müller, U. K., and Elliott, G. (2002), "Tests for Unit Roots and the Initial Condition," Discussion Paper 2002-2, University of St. Gallen, Department of Economics.

——— (2003), "Tests for Unit Roots and the Initial Condition," *Econometrica*, 71, 1269–1286.

Phillips, P. C. B. (1987), "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535–547.

Table 1. Powers of $\hat{\tau}^{\mu}_{\mathrm{AV}}(\hat{\alpha}; v, m)$, 5% Nominal Size, $T = 100$, $c = -10$.

v	m	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 3$	$\alpha = 3.5$	$\alpha = 4$
0.25	0.50	0.61	0.58	0.50	0.41	0.36	0.36	0.38	0.42	0.48
0.25	0.75	0.62	0.59	0.50	0.41	0.35	0.34	0.36	0.41	0.46
0.25	1.00	0.62	0.59	0.50	0.41	0.35	0.33	0.35	0.39	0.44
0.25	1.25	0.63	0.60	0.50	0.40	0.34	0.32	0.33	0.37	0.42
0.25	1.50	0.63	0.60	0.50	0.40	0.33	0.30	0.31	0.35	0.40
0.25	1.75	0.64	0.61	0.51	0.39	0.31	0.29	0.30	0.33	0.38
0.25	2.00	0.65	0.61	0.51	0.39	0.31	0.27	0.28	0.31	0.35
0.50	0.50	0.58	0.55	0.49	0.44	0.43	0.47	0.52	0.59	0.67
0.50	0.75	0.59	0.56	0.49	0.43	0.42	0.45	0.50	0.57	0.65
0.50	1.00	0.60	0.57	0.49	0.42	0.40	0.43	0.48	0.55	0.63
0.50	1.25	0.61	0.58	0.49	0.41	0.39	0.41	0.46	0.53	0.60
0.50	1.50	0.62	0.59	0.49	0.40	0.37	0.38	0.43	0.50	0.57
0.50	1.75	0.63	0.60	0.49	0.39	0.35	0.36	0.40	0.47	0.54
0.50	2.00	0.64	0.60	0.49	0.38	0.32	0.33	0.37	0.43	0.51
0.75	0.50	0.53	0.52	0.47	0.45	0.47	0.52	0.60	0.68	0.76
0.75	0.75	0.55	0.53	0.47	0.44	0.45	0.50	0.58	0.66	0.74
0.75	1.00	0.57	0.54	0.47	0.43	0.44	0.49	0.56	0.64	0.73
0.75	1.25	0.59	0.56	0.47	0.42	0.42	0.47	0.54	0.62	0.70
0.75	1.50	0.60	0.57	0.47	0.40	0.40	0.44	0.51	0.59	0.68
0.75	1.75	0.62	0.58	0.47	0.38	0.36	0.41	0.47	0.55	0.64
0.75	2.00	0.64	0.59	0.47	0.37	0.34	0.37	0.44	0.52	0.60
1.00	0.50	0.50	0.50	0.46	0.45	0.49	0.56	0.65	0.72	0.80
1.00	0.75	0.52	0.50	0.45	0.44	0.48	0.55	0.63	0.71	0.79
1.00	1.00	0.54	0.52	0.45	0.42	0.46	0.52	0.61	0.70	0.78
1.00	1.25	0.56	0.53	0.45	0.41	0.44	0.50	0.59	0.68	0.76
1.00	1.50	0.58	0.55	0.45	0.39	0.41	0.47	0.56	0.65	0.74
1.00	1.75	0.61	0.57	0.45	0.38	0.38	0.44	0.53	0.62	0.71
1.00	2.00	0.63	0.58	0.45	0.35	0.34	0.40	0.48	0.57	0.67
1.25	0.50	0.48	0.47	0.45	0.46	0.51	0.59	0.67	0.75	0.82
1.25	0.75	0.49	0.49	0.45	0.45	0.50	0.58	0.66	0.74	0.82
1.25	1.00	0.51	0.49	0.44	0.43	0.48	0.56	0.65	0.73	0.81
1.25	1.25	0.54	0.51	0.43	0.41	0.45	0.53	0.63	0.72	0.79
1.25	1.50	0.56	0.53	0.43	0.39	0.42	0.50	0.60	0.69	0.78
1.25	1.75	0.59	0.55	0.43	0.36	0.39	0.47	0.56	0.66	0.75
1.25	2.00	0.62	0.57	0.43	0.34	0.35	0.42	0.52	0.62	0.72
1.50	0.50	0.46	0.46	0.44	0.47	0.53	0.61	0.68	0.76	0.83
1.50	0.75	0.47	0.47	0.44	0.45	0.51	0.59	0.68	0.76	0.83
1.50	1.00	0.49	0.48	0.43	0.43	0.49	0.58	0.67	0.75	0.83
1.50	1.25	0.52	0.49	0.42	0.41	0.47	0.56	0.65	0.74	0.82
1.50	1.50	0.55	0.50	0.41	0.38	0.43	0.53	0.63	0.72	0.80
1.50	1.75	0.58	0.53	0.41	0.35	0.39	0.48	0.59	0.69	0.78
1.50	2.00	0.61	0.55	0.41	0.32	0.35	0.44	0.54	0.65	0.75

Table 2. Critical Values for $\hat{\tau}^{i}_{\text{AV}}(\hat{\alpha}; 0.73, 1.25)$.

	$\hat{\tau}^{\mu}_{\text{AV}}(\hat{\alpha}; 0.73, 1.25)$			$\hat{\tau}^{\tau}_{\text{AV}}(\hat{\alpha}; 0.73, 1.25)$		
T	10%	5%	1%	10%	5%	1%
50	-2.19	-2.51	-3.14	-3.00	-3.30	-3.92
100	-2.08	-2.38	-2.97	-2.88	-3.17	-3.73
200	-2.00	-2.30	-2.88	-2.81	-3.10	-3.64
∞	-1.91	-2.21	-2.80	-2.75	-3.03	-3.60

Table 3. Results of Tests for US PPI Inflation.

Start date	$\hat{\tau}^{\mu}_{\text{ADF}}$	$\hat{\tau}^{\mu}_{\text{ERS}}$	$\hat{\tau}^{\mu}_{\text{AV}}(\hat{\alpha}; 0.73, 1.25)$	$\hat{\alpha}$
1973:1	-2.84*	-1.26	-2.00*	1.09
1973:2	-3.24**	-0.74	-2.31**	1.97
1973:3	-3.23**	-0.57	-2.48**	2.53
1973:4	-3.09**	-2.79***	-2.88***	0.17
1973:5	-3.02**	-0.71	-2.25**	2.18
1973:6	-3.10**	-0.63	-2.37**	2.43
1973:7	-3.06**	-0.89	-2.31**	-2.12
1973:8	-3.81***	0.01	-3.77***	7.31
1973:9	-4.04***	-1.25	-3.35***	-2.79
1973:10	-3.61***	-1.50	-2.70**	-1.64
1973:11	-3.83***	-3.36***	-3.51***	0.16
1973:12	-3.44***	-0.73	-2.56**	2.24
1974:1	-3.18**	-0.40	-2.91***	4.32
1974:2	-2.99**	-0.70	-2.32**	2.48
1974:3	-2.65*	-1.17	-1.92*	1.29
1974:4	-2.42	-1.62	-1.94*	0.71
1974:5	-2.55	-0.93	-1.91	1.83
1974:6	-2.46	-2.36**	-2.39**	0.14
1974:7	-2.38	-0.56	-2.29**	5.23
1974:8	-2.47	-0.56	-2.35**	4.93
1974:9	-2.36	-1.96**	-2.05*	-0.42
1974:10	-2.38	-0.87	-1.91*	2.34
1974:11	-2.39	-1.47	-1.88	0.95
1974:12	-2.33	-1.43	-1.61	-0.69
1975:1	-2.33	-2.30**	-2.31**	-0.14
1975:2	-2.33	-1.09	-1.29	-0.96
1975:3	-2.32	-0.79	-1.00	-1.24
1975:4	-2.31	-1.34	-1.82	1.21
1975:5	-2.35	-1.75*	-1.98*	0.66
1975:6	-2.32	-2.29**	-2.30**	-0.15

NOTE: *, ** and *** denote significance at the 10%-, 5%- and 1%-levels respectively.

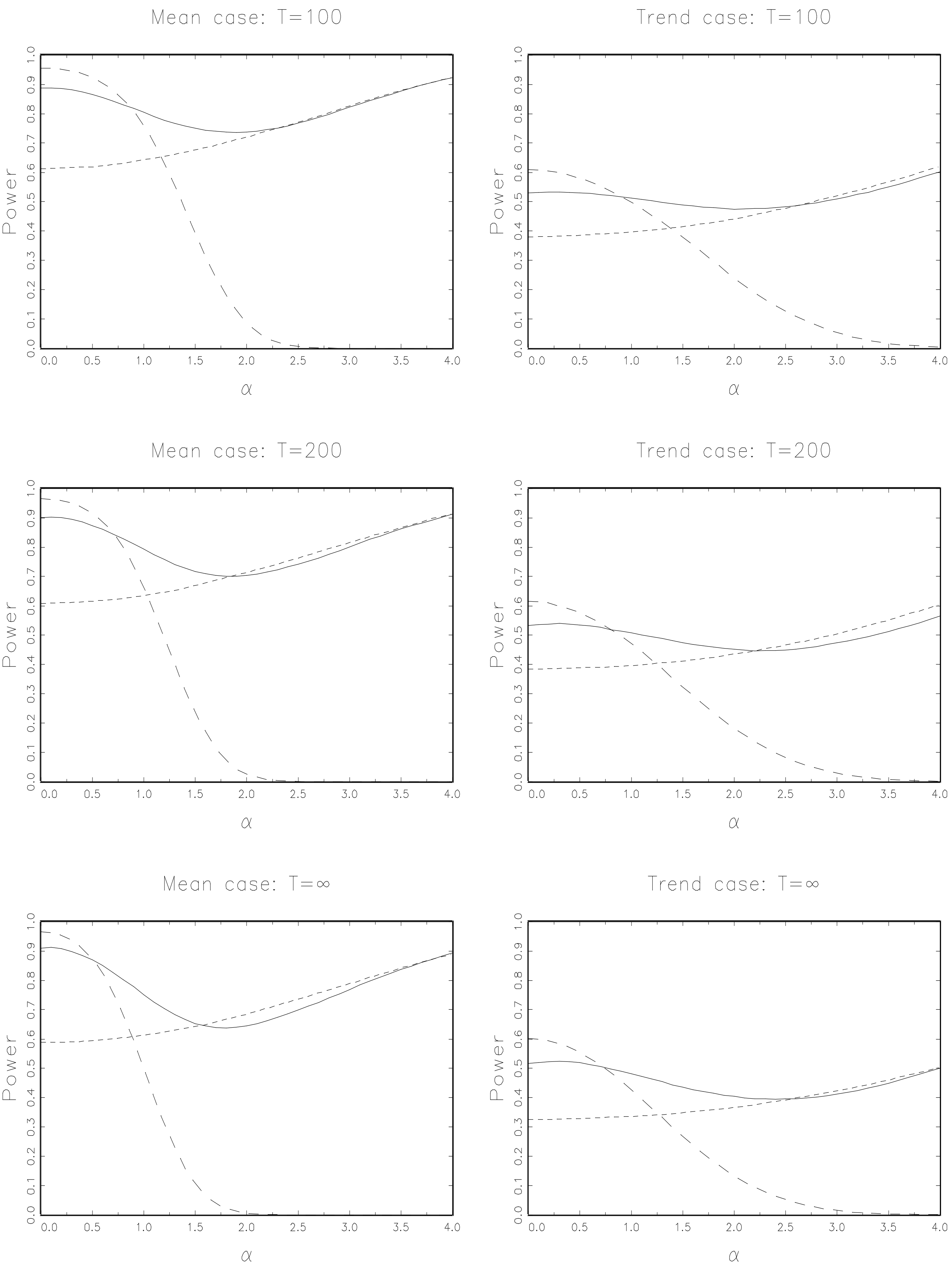

Figure 1. Powers of Tests, 5% Nominal Size, $c = -15$: - - -, $\hat{\tau}^i_{\mathrm{ADF}}$; — — , $\hat{\tau}^i_{\mathrm{ERS}}$; ——, $\hat{\tau}^i_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$.

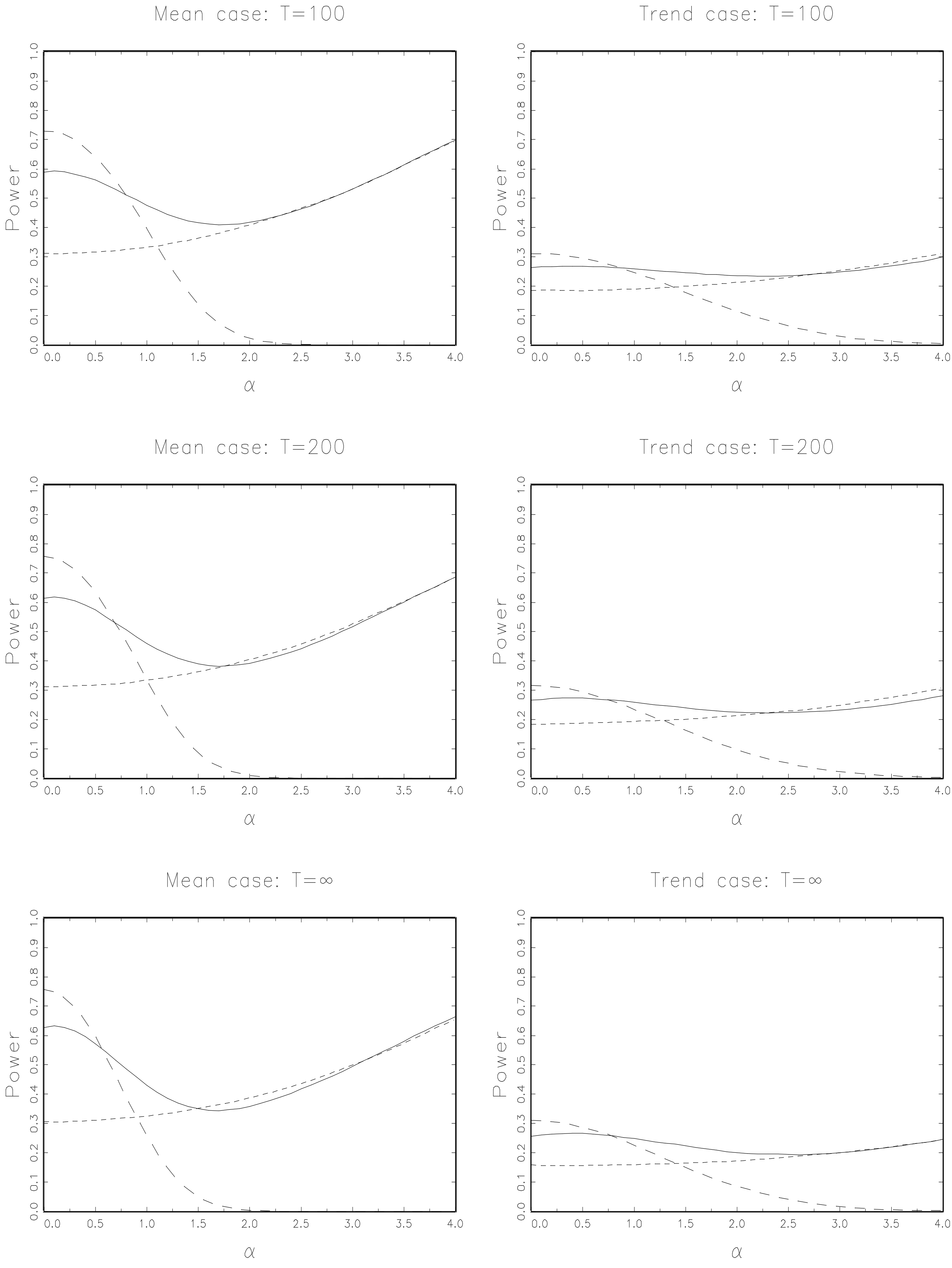

Figure 2. Powers of Tests, 5% Nominal Size, $c = -10$: - - -, $\hat{\tau}^i_{\mathrm{ADF}}$; $- -$, $\hat{\tau}^i_{\mathrm{ERS}}$; —, $\hat{\tau}^i_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$.

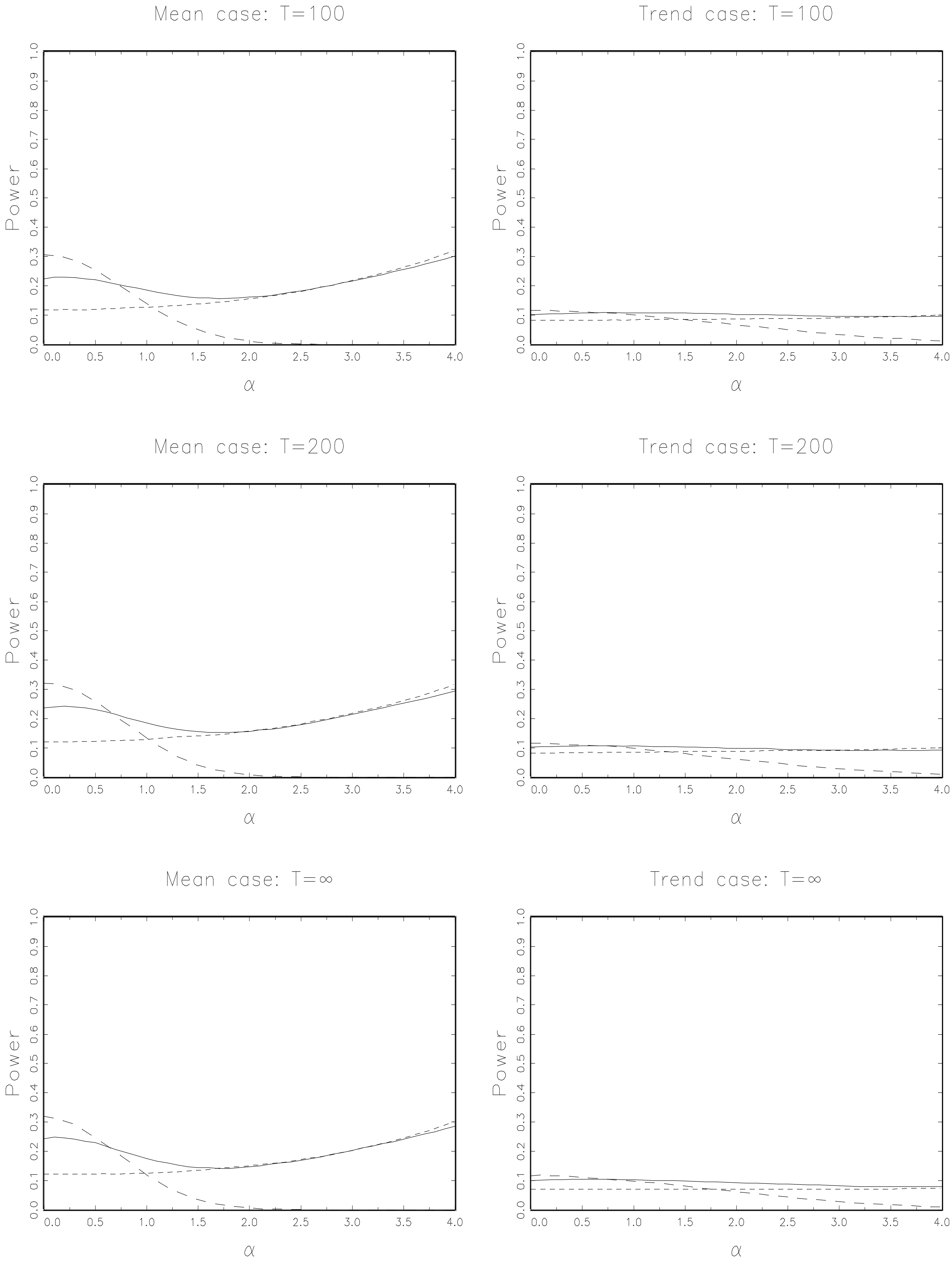

Figure 3. Powers of Tests, 5% Nominal Size, $c = -5$: - - -, $\hat{\tau}^i_{\mathrm{ADF}}$; $-\,-$, $\hat{\tau}^i_{\mathrm{ERS}}$; —, $\hat{\tau}^i_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$.

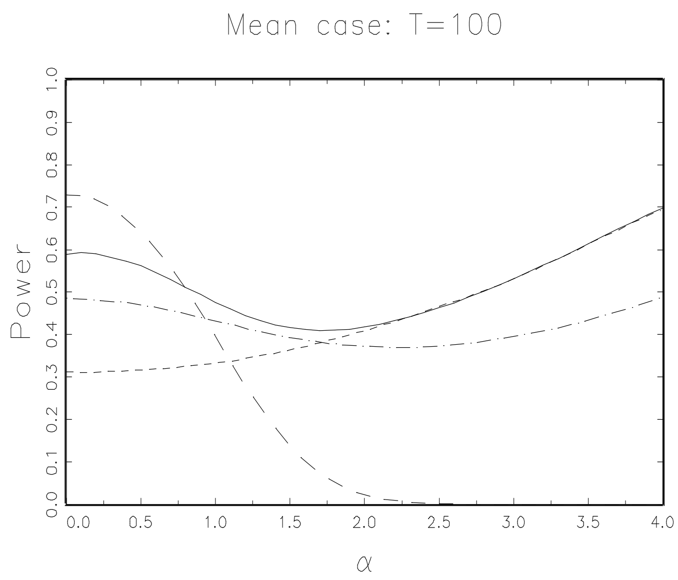

Figure 4. Powers of Tests, 5% Nominal Size, $c = -10$: - - -, $\hat{\tau}^{\mu}_{\mathrm{ADF}}$; $--$, $\hat{\tau}^{\mu}_{\mathrm{ERS}}$; —, $\hat{\tau}^{\mu}_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$; $- \cdot -$, $\hat{\tau}^{\mu}_{\mathrm{AV}}(0.75)$.

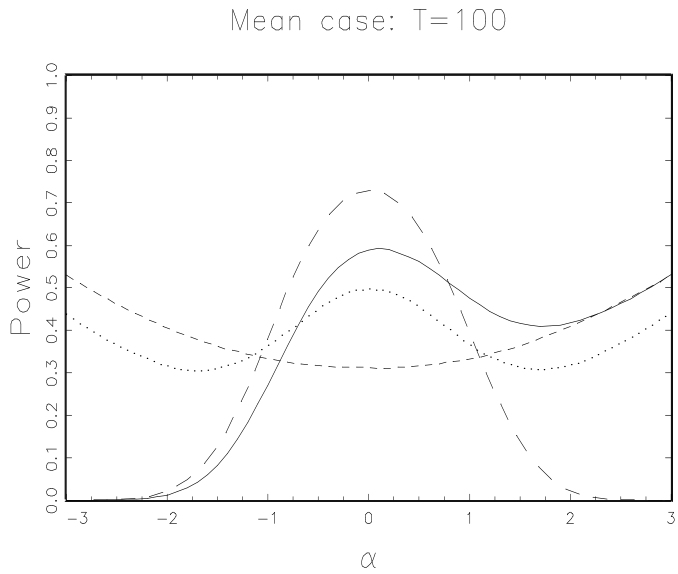

Figure 5. Powers of Tests, 5% Nominal Size, $c = -10$: - - -, $\hat{\tau}^{\mu}_{\mathrm{ADF}}$; $--$, $\hat{\tau}^{\mu}_{\mathrm{ERS}}$; —, $\hat{\tau}^{\mu}_{\mathrm{AV}}(\hat{\alpha}; 0.73, 1.25)$; $\cdots$, $\hat{\tau}^{\mu}_{\mathrm{AV}}(|\hat{\alpha}|; 0.73, 1.25)$.

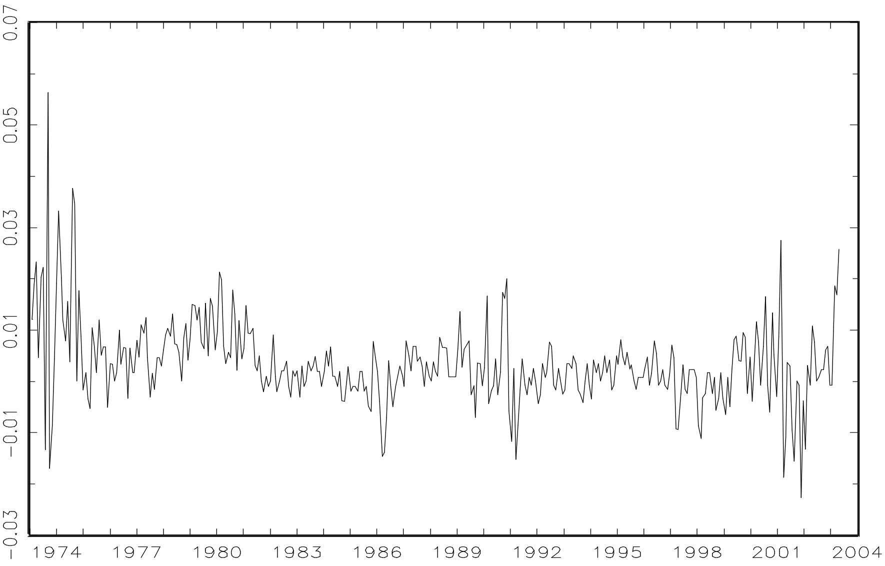

Figure 6. US PPI Inflation, 1973:1–2003:3.